

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT:
 - (A) NAME: Studiengesellschaft Kohle mbH
 - (B) STREET: Kaiser-Wilhelm-Platz 1
 - (C) CITY: Muelheim an der Ruhr
 - (E) COUNTRY: Germany
 - (F) POSTAL CODE (ZIP): 45470
- (ii) TITLE OF INVENTION: A Process for the Preperation and Identification of Novel Hydrolases Having Improved Properties
- (iii) NUMBER OF SEQUENCES: 21
- (iv) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "synthetic DNA"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GCGCAATTAA CCCTCACTAA AGGGAACAAA

30

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "synthetic DNA"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GCGTAATACG ACTCACTATA GGGCGAA

27

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1049 base pairs
 - (B) TYPE: nucleic acid

(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:85..1017

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION:163..1017

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGATCCCCCG GTTCTCCCGG AAGGATTCGG GCGATGGCTG GCAGGACGCG CCCCTCGGCC	60
CCATCAACCT GAGATGAGAA CAAC ATG AAG AAG AAG TAT CTG CTC CCC CTC	111
Met Lys Lys Lys Tyr Leu Leu Pro Leu	
-26 -25 -20	
GGC CTG GCC ATC GGT CTC GCC TCT CTC GCT GCC AGC CCT CTG ATC CAG	159
Gly Leu Ala Ile Gly Leu Ala Ser Leu Ala Ala Ser Pro Leu Ile Gln	
-15 -10 -5	
GCC AGC ACC TAC ACC CAG ACC AAA TAC CCC ATC GTG CTG GCC CAC GGC	207
Ala Ser Thr Tyr Thr Gln Thr Lys Tyr Pro Ile Val Leu Ala His Gly	
1 5 10 15	
ATG CTC GGC TTC GAC AAC ATC CTC GGG GTC GAC TAC TGG TTC GGC ATT	255
Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp Phe Gly Ile	
20 25 30	
CCC AGC GCC TTG CGC CGT GAC GGT GCC CAG GTC TAC GTC ACC GAA GTC	303
Pro Ser Ala Leu Arg Arg Asp Gly Ala Gln Val Tyr Val Thr Glu Val	
35 40 45	
AGC CAG TTG GAC ACC TCG GAA GTC CGC GGC GAG CAG TTG CTG CAA CAG	351
Ser Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gln Leu Leu Gln Gln	
50 55 60	
GTG GAG GAA ATC GTC GCC CTC AGC GGC CAG CCC AAG GTC AAC CTG ATC	399
Val Glu Glu Ile Val Ala Leu Ser Gly Gln Pro Lys Val Asn Leu Ile	
65 70 75	
GGC CAC AGC CAC GGC GGG CCG ACC ATC CGC TAC GTC GCC GCC GTA CGT	447
Gly His Ser His Gly Gly Pro Thr Ile Arg Tyr Val Ala Ala Val Arg	
80 85 90 95	
CCC GAC CTG ATC GCT TCC GCC ATC AGC GTC GGC GCC CCG CAC AAG GGT	495
Pro Asp Leu Ile Ala Ser Ala Ile Ser Val Gly Ala Pro His Lys Gly	
100 105 110	
TCG GAC ACC GCC GAC TTC CTG CGC CAG ATC CCA CCG GGT TCG GCC GGC	543
Ser Asp Thr Ala Asp Phe Leu Arg Gln Ile Pro Pro Gly Ser Ala Gly	
115 120 125	
GAG GCA GTC CTC TCC GGG CTG GTC AAC AGC CTC GGC GCG CTG ATC AGC	591
Glu Ala Val Leu Ser Gly Leu Val Asn Ser Leu Gly Ala Leu Ile Ser	
130 135 140	

TTC	CTT	TCC	AGC	GGC	GGC	ACC	GGT	ACG	CAG	AAT	TCA	CTG	GGC	TCG	CTG	639
Phe	Leu	Ser	Ser	Gly	Gly	Thr	Gly	Thr	Gln	Asn	Ser	Leu	Gly	Ser	Leu	
	145					150					155					
GAG	TCG	CTG	AAC	AGC	GAG	GGT	GCC	GCG	CGC	TTC	AAC	GCC	AAG	TAC	CCG	687
Glu	Ser	Leu	Asn	Ser	Glu	Gly	Ala	Ala	Arg	Phe	Asn	Ala	Lys	Tyr	Pro	
160					165					170					175	
CAG	GGC	ATC	CCC	ACC	TCG	GCC	TGC	GGC	GAA	GGC	GCC	TAC	AAG	GTC	AAC	735
Gln	Gly	Ile	Pro	Thr	Ser	Ala	Cys	Gly	Glu	Gly	Ala	Tyr	Lys	Val	Asn	
				180					185					190		
GGC	GTG	AGC	TAT	TAC	TCC	TGG	AGC	GGT	TCC	TCG	CCG	CTG	ACC	AAC	TTC	783
Gly	Val	Ser	Tyr	Tyr	Ser	Trp	Ser	Gly	Ser	Ser	Pro	Leu	Thr	Asn	Phe	
			195					200					205			
CTC	GAT	CCG	AGC	GAC	GCC	TTC	CTC	GGC	GCC	TCG	TCG	CTG	ACC	TTC	AAG	831
Leu	Asp	Pro	Ser	Asp	Ala	Phe	Leu	Gly	Ala	Ser	Ser	Leu	Thr	Phe	Lys	
		210					215					220				
AAC	GGC	ACC	GCC	AAC	GAC	GGC	CTG	GTC	GGC	ACC	TGC	AGT	TCG	CAC	CTG	879
Asn	Gly	Thr	Ala	Asn	Asp	Gly	Leu	Val	Gly	Thr	Cys	Ser	Ser	His	Leu	
	225					230					235					
GGC	ATG	GTG	ATC	CGC	GAC	AAC	TAC	CGG	ATG	AAC	CAC	CTG	GAC	GAG	GTG	927
Gly	Met	Val	Ile	Arg	Asp	Asn	Tyr	Arg	Met	Asn	His	Leu	Asp	Glu	Val	
240					245					250				255		
AAC	CAG	GTC	TTC	GGC	CTC	ACC	AGC	CTG	TTC	GAG	ACC	AGC	CCG	GTC	AGC	975
Asn	Gln	Val	Phe	Gly	Leu	Thr	Ser	Leu	Phe	Glu	Thr	Ser	Pro	Val	Ser	
				260					265					270		
GTC	TAC	CGC	CAG	CAC	GCC	AAC	CGC	CTG	AAG	AAC	GCC	AGC	CTG			1017
Val	Tyr	Arg	Gln	His	Ala	Asn	Arg	Leu	Lys	Asn	Ala	Ser	Leu			
			275					280					285			
TAGGACCCCG GCCGGGGCCT CGGCCCGGGC CC																1049

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 311 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met	Lys	Lys	Lys	Tyr	Leu	Leu	Pro	Leu	Gly	Leu	Ala	Ile	Gly	Leu	Ala	
-26	-25					-20					-15					
Ser	Leu	Ala	Ala	Ser	Pro	Leu	Ile	Gln	Ala	Ser	Thr	Tyr	Thr	Gln	Thr	
-10				-5						1				5		
Lys	Tyr	Pro	Ile	Val	Leu	Ala	His	Gly	Met	Leu	Gly	Phe	Asp	Asn	Ile	
			10					15					20			
Leu	Gly	Val	Asp	Tyr	Trp	Phe	Gly	Ile	Pro	Ser	Ala	Leu	Arg	Arg	Asp	
	25						30					35				

Gly	Ala	Gln	Val	Tyr	Val	Thr	Glu	Val	Ser	Gln	Leu	Asp	Thr	Ser	Glu	40	45	50
Val	Arg	Gly	Glu	Gln	Leu	Leu	Gln	Gln	Val	Glu	Glu	Ile	Val	Ala	Leu	55	60	65
Ser	Gly	Gln	Pro	Lys	Val	Asn	Leu	Ile	Gly	His	Ser	His	Gly	Gly	Pro	75	80	85
Thr	Ile	Arg	Tyr	Val	Ala	Ala	Val	Arg	Pro	Asp	Leu	Ile	Ala	Ser	Ala	90	95	100
Ile	Ser	Val	Gly	Ala	Pro	His	Lys	Gly	Ser	Asp	Thr	Ala	Asp	Phe	Leu	105	110	115
Arg	Gln	Ile	Pro	Pro	Gly	Ser	Ala	Gly	Glu	Ala	Val	Leu	Ser	Gly	Leu	120	125	130
Val	Asn	Ser	Leu	Gly	Ala	Leu	Ile	Ser	Phe	Leu	Ser	Ser	Gly	Gly	Thr	135	140	145
Gly	Thr	Gln	Asn	Ser	Leu	Gly	Ser	Leu	Glu	Ser	Leu	Asn	Ser	Glu	Gly	155	160	165
Ala	Ala	Arg	Phe	Asn	Ala	Lys	Tyr	Pro	Gln	Gly	Ile	Pro	Thr	Ser	Ala	170	175	180
Cys	Gly	Glu	Gly	Ala	Tyr	Lys	Val	Asn	Gly	Val	Ser	Tyr	Tyr	Ser	Trp	185	190	195
Ser	Gly	Ser	Ser	Pro	Leu	Thr	Asn	Phe	Leu	Asp	Pro	Ser	Asp	Ala	Phe	200	205	210
Leu	Gly	Ala	Ser	Ser	Leu	Thr	Phe	Lys	Asn	Gly	Thr	Ala	Asn	Asp	Gly	215	220	225
Leu	Val	Gly	Thr	Cys	Ser	Ser	His	Leu	Gly	Met	Val	Ile	Arg	Asp	Asn	235	240	245
Tyr	Arg	Met	Asn	His	Leu	Asp	Glu	Val	Asn	Gln	Val	Phe	Gly	Leu	Thr	250	255	260
Ser	Leu	Phe	Glu	Thr	Ser	Pro	Val	Ser	Val	Tyr	Arg	Gln	His	Ala	Asn	265	270	275
Arg	Leu	Lys	Asn	Ala	Ser	Leu										280	285	

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1049 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 85..1017

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION:163..1017

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGATCCCCCG GTTCTCCCGG AAGGATTCGG GCGATGGCTG GCAGGACGCG CCCCTCGGCC	60
CCATCAACCT GAGATGAGAA CAAC ATG AAG AAG AAG TCT CTG CTC CCC CTC	111
Met Lys Lys Lys Ser Leu Leu Pro Leu	
-26 -25 -20	
GGC CTG GCC ATC GGT CTC GCC TCT CTC GCT GCC AGC CCT CTG ATC CAG	159
Gly Leu Ala Ile Gly Leu Ala Ser Leu Ala Ala Ser Pro Leu Ile Gln	
-15 -10 -5	
GCC AGC ACC TAC ACC CAG ACC AAA TAC CCC ATC GTG CTG GCC CAC GGC	207
Ala Ser Thr Tyr Thr Gln Thr Lys Tyr Pro Ile Val Leu Ala His Gly	
1 5 10 15	
ATG CTC GGC TTC GAC AAC ATC CTC GGG GTC GAC TAC TGG TTC GGC ATT	255
Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp Phe Gly Ile	
20 25 30	
CCC AGC GCC TTG CGC CGT GAC GGT GCC CAG GTC TAC GTC ACC GAA GTC	303
Pro Ser Ala Leu Arg Arg Asp Gly Ala Gln Val Tyr Val Thr Glu Val	
35 40 45	
AGC CAG TTG GAC ACC TCG GAA GTC CGC GGC GAG CAG TTG CTG CAA CAG	351
Ser Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gln Leu Gln Gln	
50 55 60	
GTG GAG GAA ATC GTC GCC CTC AGC GGC CAG CCC AAG GTC AAC CTG ATC	399
Val Glu Glu Ile Val Ala Leu Ser Gly Gln Pro Lys Val Asn Leu Ile	
65 70 75	
GGC CAC AGC CAC GGC GGG CCG ACC ATC CGC TAC GTC GCC GCC GTA CGT	447
Gly His Ser His Gly Gly Pro Thr Ile Arg Tyr Val Ala Ala Val Arg	
80 85 90 95	
CCC GAC CTG ATC GCT TCC GCC ACC AGC GTC GGC GCC CCG CAC AAG GGT	495
Pro Asp Leu Ile Ala Ser Ala Thr Ser Val Gly Ala Pro His Lys Gly	
100 105 110	
TCG GAC ACC GCC GAC TTC CTG CGC CAG ATC CCA CCG GGT TCG GCC GGC	543
Ser Asp Thr Ala Asp Phe Leu Arg Gln Ile Pro Pro Gly Ser Ala Gly	
115 120 125	
GAG GCA GTC CTC TCC GGG CTG GTC AAC AGC CTC GGC GCG CTG ATC AGC	591
Glu Ala Val Leu Ser Gly Leu Val Asn Ser Leu Gly Ala Leu Ile Ser	
130 135 140	
TTC CTT TCC AGC GGC GGC ACC GGT ACG CAG AAT TCA CTG GGC TCG CTG	639
Phe Leu Ser Ser Gly Gly Thr Gly Thr Gln Asn Ser Leu Gly Ser Leu	
145 150 155	
GAG TCG CTG AAC AGC GAG GGT GCC GCG CGC TTC AAC GCC AAG TAC CCG	687
Glu Ser Leu Asn Ser Glu Gly Ala Ala Arg Phe Asn Ala Lys Tyr Pro	
160 165 170 175	

CAG GGC ATC CCC ACC TCG GCC TGC GGC GAA GGC GCC TAC AAG GTC AAC	735
Gln Gly Ile Pro Thr Ser Ala Cys Gly Glu Gly Ala Tyr Lys Val Asn	
180 185 190	
GGC GTG AGC TAT TAC TCC TGG AGC GGT TCC TCG CCG CTG ACC AAC TTC	783
Gly Val Ser Tyr Tyr Ser Trp Ser Gly Ser Ser Pro Leu Thr Asn Phe	
195 200 205	
CTC GAT CCG AGC GAC GCC TTC CTC GGC GCC TCG TCG CTG ACC TTC AAG	831
Leu Asp Pro Ser Asp Ala Phe Leu Gly Ala Ser Ser Leu Thr Phe Lys	
210 215 220	
AAC GGC ACC GCC AAC GAC GGC CTG GTC GGC ACC TGC AGT TCG CAC CTG	879
Asn Gly Thr Ala Asn Asp Gly Leu Val Gly Thr Cys Ser Ser His Leu	
225 230 235	
GGC ATG GTG ATC CGC GAC AAC TAC CGG ATG AAC CAC CTG GAC GAG GTG	927
Gly Met Val Ile Arg Asp Asn Tyr Arg Met Asn His Leu Asp Glu Val	
240 245 250 255	
AAC CAG GTC TTC GGC CTC ACC AGC CTG TTC GAG ACC AGC CCG GTC AGC	975
Asn Gln Val Phe Gly Leu Thr Ser Leu Phe Glu Thr Ser Pro Val Ser	
260 265 270	
GTC TAC CGC CAG CAC GCC AAC CGC CTG AAG AAC GCC AGC CTG	1017
Val Tyr Arg Gln His Ala Asn Arg Leu Lys Asn Ala Ser Leu	
275 280 285	
TAGGACCCCG GCCGGGGCCT CGGCCCGGGC CC	1049

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 311 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala	
-26 -25 -20 -15	
Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr	
-10 -5 1 5	
Lys Tyr Pro Ile Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile	
10 15 20	
Leu Gly Val Asp Tyr Trp Phe Gly Ile Pro Ser Ala Leu Arg Arg Asp	
25 30 35	
Gly Ala Gln Val Tyr Val Thr Glu Val Ser Gln Leu Asp Thr Ser Glu	
40 45 50	
Val Arg Gly Glu Gln Leu Leu Gln Gln Val Glu Glu Ile Val Ala Leu	
55 60 65 70	
Ser Gly Gln Pro Lys Val Asn Leu Ile Gly His Ser His Gly Gly Pro	
75 80 85	

Thr Ile Arg Tyr Val Ala Ala Val Arg Pro Asp Leu Ile Ala Ser Ala
90 95 100

Thr Ser Val Gly Ala Pro His Lys Gly Ser Asp Thr Ala Asp Phe Leu
105 110 115

Arg Gln Ile Pro Pro Gly Ser Ala Gly Glu Ala Val Leu Ser Gly Leu
120 125 130

Val Asn Ser Leu Gly Ala Leu Ile Ser Phe Leu Ser Ser Gly Gly Thr
135 140 145 150

Gly Thr Gln Asn Ser Leu Gly Ser Leu Glu Ser Leu Asn Ser Glu Gly
155 160 165

Ala Ala Arg Phe Asn Ala Lys Tyr Pro Gln Gly Ile Pro Thr Ser Ala
170 175 180

Cys Gly Glu Gly Ala Tyr Lys Val Asn Gly Val Ser Tyr Tyr Ser Trp
185 190 195

Ser Gly Ser Ser Pro Leu Thr Asn Phe Leu Asp Pro Ser Asp Ala Phe
200 205 210

Leu Gly Ala Ser Ser Leu Thr Phe Lys Asn Gly Thr Ala Asn Asp Gly
215 220 225 230

Leu Val Gly Thr Cys Ser Ser His Leu Gly Met Val Ile Arg Asp Asn
235 240 245

Tyr Arg Met Asn His Leu Asp Glu Val Asn Gln Val Phe Gly Leu Thr
250 255 260

Ser Leu Phe Glu Thr Ser Pro Val Ser Val Tyr Arg Gln His Ala Asn
265 270 275

Arg Leu Lys Asn Ala Ser Leu
280 285

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1049 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 85..1017

- (ix) FEATURE:
 - (A) NAME/KEY: mat_peptide
 - (B) LOCATION: 163..1017

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGATCCCCCG GTTCTCCCGG AAGGATTCGG GCGATGGCTG GCAGGACGCG CCCCTCGGCC 60

CCATCAACCT	GAGATGAGAA	CAAC	ATG	AAG	AAG	AAG	TCT	CTG	CTC	CCC	CTC					111
			Met	Lys	Lys	Lys	Ser	Leu	Leu	Pro	Leu					
			-26	-25						-20						
GGC	CTG	GCC	ATC	GGT	CTC	GCC	TCT	CTC	GCT	GCC	AGC	CCT	CTG	ATC	CAG	159
Gly	Leu	Ala	Ile	Gly	Leu	Ala	Ser	Leu	Ala	Ala	Ser	Pro	Leu	Ile	Gln	
		-15					-10					-5				
GCC	AGC	ACC	TAC	ACC	CAG	ACC	AAA	TAC	CCC	ATC	GTG	CTG	GCC	CAC	GGC	207
Ala	Ser	Thr	Tyr	Thr	Gln	Thr	Lys	Tyr	Pro	Ile	Val	Leu	Ala	His	Gly	
	1				5					10					15	
ATG	CTC	GGC	TTC	GAC	AAC	ATC	CTT	GGG	GTC	GAC	TAC	TGG	TTC	GGC	ATT	255
Met	Leu	Gly	Phe	Asp	Asn	Ile	Leu	Gly	Val	Asp	Tyr	Trp	Phe	Gly	Ile	
				20					25					30		
CCC	AGC	GCC	TTG	CGC	CGT	GAC	GGT	GCC	CAG	GTC	TAC	GTC	ACC	GAA	GTC	303
Pro	Ser	Ala	Leu	Arg	Arg	Asp	Gly	Ala	Gln	Val	Tyr	Val	Thr	Glu	Val	
			35					40					45			
AGC	CAG	TTG	GAC	ACC	TCG	GAA	GTC	CGC	GGC	GAG	CAG	TTG	CTG	CAA	CAG	351
Ser	Gln	Leu	Asp	Thr	Ser	Glu	Val	Arg	Gly	Glu	Gln	Leu	Leu	Gln	Gln	
		50					55					60				
GTG	GAG	GAA	ATC	GTC	GCC	CTC	AGC	GGC	CAG	CCC	AAG	GTC	AAC	CTG	ATC	399
Val	Glu	Glu	Ile	Val	Ala	Leu	Ser	Gly	Gln	Pro	Lys	Val	Asn	Leu	Ile	
	65					70					75					
GGC	CAC	AGC	CAC	GGC	GGG	CCG	ACC	ATC	CGC	TAC	GTC	GCC	GCC	GTA	CGT	447
Gly	His	Ser	His	Gly	Gly	Pro	Thr	Ile	Arg	Tyr	Val	Ala	Ala	Val	Arg	
	80				85				90						95	
CCC	GAC	CTG	ATC	GCT	TCC	GCC	ACC	AGC	GTC	GGC	GCC	CCG	CAC	AAG	GGT	495
Pro	Asp	Leu	Ile	Ala	Ser	Ala	Thr	Ser	Val	Gly	Ala	Pro	His	Lys	Gly	
				100					105					110		
TCG	GAC	ACC	GCC	GAC	TTC	CTG	CGC	CAG	ATC	CCA	CCG	GGT	TCG	GCC	GGC	543
Ser	Asp	Thr	Ala	Asp	Phe	Leu	Arg	Gln	Ile	Pro	Pro	Gly	Ser	Ala	Gly	
			115					120					125			
GAG	GCA	GTC	CTC	TCC	GGG	CTG	GTC	AAC	AGC	CTC	GGC	GCG	CTG	ATC	AGC	591
Glu	Ala	Val	Leu	Ser	Gly	Leu	Val	Asn	Ser	Leu	Gly	Ala	Leu	Ile	Ser	
		130					135					140				
TTC	CTT	TCC	AGC	GGC	GGC	ACC	GGT	ACG	CAG	AAT	TTA	CTG	GGC	TCG	CTG	639
Phe	Leu	Ser	Ser	Gly	Gly	Thr	Gly	Thr	Gln	Asn	Leu	Leu	Gly	Ser	Leu	
	145					150						155				
GAG	TCG	CTG	AAC	AGC	GAG	GGT	GCC	GCG	CGC	TTC	AAC	GCC	AAG	TAC	CCG	687
Glu	Ser	Leu	Asn	Ser	Glu	Gly	Ala	Ala	Arg	Phe	Asn	Ala	Lys	Tyr	Pro	
	160				165					170					175	
CAG	GGC	ATC	CCC	ACC	TCG	GCC	TGC	GGC	GAA	GGC	GCC	TAC	AAG	GTC	AAC	735
Gln	Gly	Ile	Pro	Thr	Ser	Ala	Cys	Gly	Glu	Gly	Ala	Tyr	Lys	Val	Asn	
				180					185					190		
GGC	GTG	AGC	TAT	TAC	TCC	TGG	AGC	GGT	TCC	TCG	CCG	CTG	ACC	AAC	TTC	783
Gly	Val	Ser	Tyr	Tyr	Ser	Trp	Ser	Gly	Ser	Ser	Pro	Leu	Thr	Asn	Phe	
			195					200					205			

CTC GAT CCG AGC GAC GCC TTC CTC GGC GCC TCG TCG CTG ACC TTC AAG	831
Leu Asp Pro Ser Asp Ala Phe Leu Gly Ala Ser Ser Leu Thr Phe Lys	
210 215 220	
AAC GGC ACC GCC AAC GAC GGC CTG GTC GGC ACC TGC AGT TCG CAC CTG	879
Asn Gly Thr Ala Asn Asp Gly Leu Val Gly Thr Cys Ser Ser His Leu	
225 230 235	
GGC ATG GTG ATC CGC GAC AAC TAC CGG ATG AAC CAC CTG GAC GAG GTG	927
Gly Met Val Ile Arg Asp Asn Tyr Arg Met Asn His Leu Asp Glu Val	
240 245 250 255	
AAC CAG GTC TTC GGC CTC ACC AGC CTG TTC GAG ACC AGC CCG GTC AGC	975
Asn Gln Val Phe Gly Leu Thr Ser Leu Phe Glu Thr Ser Pro Val Ser	
260 265 270	
GTC TAC CGC CAG CAC GCC AAC CGC CTG AAG AAC GCC AGC CTG	1017
Val Tyr Arg Gln His Ala Asn Arg Leu Lys Asn Ala Ser Leu	
275 280 285	
TAGGACCCCG GCCGGGGCCT CGGCCCCGGGC CC	1049

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 311 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala	
-26 -25 -20 -15	
Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr	
-10 -5 1 5	
Lys Tyr Pro Ile Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile	
10 15 20	
Leu Gly Val Asp Tyr Trp Phe Gly Ile Pro Ser Ala Leu Arg Arg Asp	
25 30 35	
Gly Ala Gln Val Tyr Val Thr Glu Val Ser Gln Leu Asp Thr Ser Glu	
40 45 50	
Val Arg Gly Glu Gln Leu Leu Gln Gln Val Glu Glu Ile Val Ala Leu	
55 60 65 70	
Ser Gly Gln Pro Lys Val Asn Leu Ile Gly His Ser His Gly Gly Pro	
75 80 85	
Thr Ile Arg Tyr Val Ala Ala Val Arg Pro Asp Leu Ile Ala Ser Ala	
90 95 100	
Thr Ser Val Gly Ala Pro His Lys Gly Ser Asp Thr Ala Asp Phe Leu	
105 110 115	
Arg Gln Ile Pro Pro Gly Ser Ala Gly Glu Ala Val Leu Ser Gly Leu	
120 125 130	

Val	Asn	Ser	Leu	Gly	Ala	Leu	Ile	Ser	Phe	Leu	Ser	Ser	Gly	Gly	Thr	
135					140					145					150	
Gly	Thr	Gln	Asn	Leu	Leu	Gly	Ser	Leu	Glu	Ser	Leu	Asn	Ser	Glu	Gly	
				155					160					165		
Ala	Ala	Arg	Phe	Asn	Ala	Lys	Tyr	Pro	Gln	Gly	Ile	Pro	Thr	Ser	Ala	
			170					175					180			
Cys	Gly	Glu	Gly	Ala	Tyr	Lys	Val	Asn	Gly	Val	Ser	Tyr	Tyr	Ser	Trp	
		185					190					195				
Ser	Gly	Ser	Ser	Pro	Leu	Thr	Asn	Phe	Leu	Asp	Pro	Ser	Asp	Ala	Phe	
	200					205					210					
Leu	Gly	Ala	Ser	Ser	Leu	Thr	Phe	Lys	Asn	Gly	Thr	Ala	Asn	Asp	Gly	
215					220					225					230	
Leu	Val	Gly	Thr	Cys	Ser	Ser	His	Leu	Gly	Met	Val	Ile	Arg	Asp	Asn	
				235					240					245		
Tyr	Arg	Met	Asn	His	Leu	Asp	Glu	Val	Asn	Gln	Val	Phe	Gly	Leu	Thr	
			250				255						260			
Ser	Leu	Phe	Glu	Thr	Ser	Pro	Val	Ser	Val	Tyr	Arg	Gln	His	Ala	Asn	
		265					270					275				
Arg	Leu	Lys	Asn	Ala	Ser	Leu										
280						285										

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1047 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 84..1016

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 162..1016

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGATCCCCGG	TTCTCCCGGA	AGGATTCGGG	CGATGGCTGG	CAGGACGCGC	CCCTCGGCCC	60
CATCAACCTG	AGATGAGAAC	AAC ATG AAG AAG	TCT CTG CTC	CCC CTC		110
		Met Lys Lys Lys	Ser Leu Leu	Pro Leu		
		-26 -25		-20		
GGC CTG GCC	ATC GGT CTC	GCC TCT CTC	GCT GCC AGC	CCT CTG	ATC CAG	158
Gly Leu Ala	Ile Gly Leu	Ala Ser Leu	Ala Ala Ser	Pro Leu	Ile Gln	
	-15		-10		-5	

GCC	AGC	ACC	TAC	ACC	CAG	ACC	AAA	TAC	CCC	ATC	GTG	CTG	GCC	CAC	GGC	206
Ala	Ser	Thr	Tyr	Thr	Gln	Thr	Lys	Tyr	Pro	Ile	Val	Leu	Ala	His	Gly	
	1				5					10					15	
ATG	CTC	GGC	TTC	GAC	AAC	ATC	CTC	GGG	GTC	GAC	TAC	TGG	TTC	GGC	ATT	254
Met	Leu	Gly	Phe	Asp	Asn	Ile	Leu	Gly	Val	Asp	Tyr	Trp	Phe	Gly	Ile	
				20					25					30		
CCC	AGC	GCC	TTG	CGC	CGT	GAC	GGT	GCC	CAG	GTC	TAC	GTC	ACC	GAA	GTC	302
Pro	Ser	Ala	Leu	Arg	Arg	Asp	Gly	Ala	Gln	Val	Tyr	Val	Thr	Glu	Val	
			35					40					45			
AGC	CAG	TTG	GAC	ACC	TCG	GAA	GTC	CGC	GGC	GAG	CAG	TTG	CTG	CAA	CAG	350
Ser	Gln	Leu	Asp	Thr	Ser	Glu	Val	Arg	Gly	Glu	Gln	Leu	Leu	Gln	Gln	
		50					55					60				
GTG	GAG	GAA	ATC	GTC	GCC	CTC	AGC	GGC	CAG	CCC	AAG	GTC	AAC	CTG	ATC	398
Val	Glu	Glu	Ile	Val	Ala	Leu	Ser	Gly	Gln	Pro	Lys	Val	Asn	Leu	Ile	
	65					70					75					
GGC	CAC	AGC	CAC	GGC	GGG	CCG	ACC	ATC	CGC	TAC	GTC	GCC	GCC	GTA	CGT	446
Gly	His	Ser	His	Gly	Gly	Pro	Thr	Ile	Arg	Tyr	Val	Ala	Ala	Val	Arg	
	80				85					90					95	
CCC	GAC	CTG	ATC	GCT	TCC	GCC	ACC	AGC	GTC	GGC	GCC	CCG	CAC	AAG	GGT	494
Pro	Asp	Leu	Ile	Ala	Ser	Ala	Thr	Ser	Val	Gly	Ala	Pro	His	Lys	Gly	
				100					105					110		
TCG	GAC	ACC	GCC	GAC	TTC	CTG	CGC	CAG	ATC	CCA	CCG	GGT	TCG	GCC	GGC	542
Ser	Asp	Thr	Ala	Asp	Phe	Leu	Arg	Gln	Ile	Pro	Pro	Gly	Ser	Ala	Gly	
			115					120					125			
GAG	GCA	GTC	CTC	TCC	GGG	CTG	GTC	AAC	AGC	CTC	GGC	GCG	CTG	ATC	AGC	590
Glu	Ala	Val	Leu	Ser	Gly	Leu	Val	Asn	Ser	Leu	Gly	Ala	Leu	Ile	Ser	
		130					135					140				
TTC	CTT	TCC	AGC	GGC	AGC	ACC	GGT	ACG	CAG	AAT	TCA	CTG	GGC	TCG	CTG	638
Phe	Leu	Ser	Ser	Gly	Ser	Thr	Gly	Thr	Gln	Asn	Ser	Leu	Gly	Ser	Leu	
	145					150					155					
GAG	TCG	CTG	AAC	AGC	GAG	GGT	GCC	GCG	CGC	TTC	AAC	GCC	AAG	TAC	CCG	686
Glu	Ser	Leu	Asn	Ser	Glu	Gly	Ala	Ala	Arg	Phe	Asn	Ala	Lys	Tyr	Pro	
	160				165					170					175	
CAG	GGC	ATC	CCC	ACC	TCG	GCC	TGC	GGC	GAA	GGC	GCC	TAC	AAG	GTC	AAC	734
Gln	Gly	Ile	Pro	Thr	Ser	Ala	Cys	Gly	Glu	Gly	Ala	Tyr	Lys	Val	Asn	
				180					185					190		
GGC	GTG	AGC	TAT	TAC	TCC	TGG	AGC	GGT	TCC	TCG	CCG	CTG	ACC	AAC	TTC	782
Gly	Val	Ser	Tyr	Tyr	Ser	Trp	Ser	Gly	Ser	Ser	Pro	Leu	Thr	Asn	Phe	
			195					200					205			
CTC	GAT	CCG	AGC	GAC	GCC	TTC	CTC	GGC	GCC	TCG	TCG	CTG	ACC	TTC	AAG	830
Leu	Asp	Pro	Ser	Asp	Ala	Phe	Leu	Gly	Ala	Ser	Ser	Leu	Thr	Phe	Lys	
		210					215					220				
AAC	GGC	ACC	GCC	AAC	GAC	GGC	CTG	GTC	GGC	ACC	TGC	AGT	TCG	CAC	CTG	878
Asn	Gly	Thr	Ala	Asn	Asp	Gly	Leu	Val	Gly	Thr	Cys	Ser	Ser	His	Leu	
	225					230					235					

GGC ATG GTG ATC CGC GAC AAC TAC CGG ATG AAC CAC CTG GAC GAG GTG	926
Gly Met Val Ile Arg Asp Asn Tyr Arg Met Asn His Leu Asp Glu Val	
240 245 250 255	
AAC CAG GTC TTC GGC CTC ACC AGC CTG TTC GAG ACC AGC CCG GTC AGC	974
Asn Gln Val Phe Gly Leu Thr Ser Leu Phe Glu Thr Ser Pro Val Ser	
260 265 270	
GTC TAC CGC CAG CAC GCC AAC CGC CTG AAG AAC GCC AGC CTG	1016
Val Tyr Arg Gln His Ala Asn Arg Leu Lys Asn Ala Ser Leu	
275 280 285	
TAGGACCCCG GCCGGGGCCT CGGCCCGGGC C	1047

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 311 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala	
-26 -25 -20 -15	
Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr	
-10 -5 1 5	
Lys Tyr Pro Ile Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile	
10 15 20	
Leu Gly Val Asp Tyr Trp Phe Gly Ile Pro Ser Ala Leu Arg Arg Asp	
25 30 35	
Gly Ala Gln Val Tyr Val Thr Glu Val Ser Gln Leu Asp Thr Ser Glu	
40 45 50	
Val Arg Gly Glu Gln Leu Leu Gln Gln Val Glu Glu Ile Val Ala Leu	
55 60 65 70	
Ser Gly Gln Pro Lys Val Asn Leu Ile Gly His Ser His Gly Gly Pro	
75 80 85	
Thr Ile Arg Tyr Val Ala Ala Val Arg Pro Asp Leu Ile Ala Ser Ala	
90 95 100	
Thr Ser Val Gly Ala Pro His Lys Gly Ser Asp Thr Ala Asp Phe Leu	
105 110 115	
Arg Gln Ile Pro Pro Gly Ser Ala Gly Glu Ala Val Leu Ser Gly Leu	
120 125 130	
Val Asn Ser Leu Gly Ala Leu Ile Ser Phe Leu Ser Ser Gly Ser Thr	
135 140 145 150	
Gly Thr Gln Asn Ser Leu Gly Ser Leu Glu Ser Leu Asn Ser Glu Gly	
155 160 165	

Ala	Ala	Arg	Phe	Asn	Ala	Lys	Tyr	Pro	Gln	Gly	Ile	Pro	Thr	Ser	Ala
			170					175					180		
Cys	Gly	Glu	Gly	Ala	Tyr	Lys	Val	Asn	Gly	Val	Ser	Tyr	Tyr	Ser	Trp
		185					190					195			
Ser	Gly	Ser	Ser	Pro	Leu	Thr	Asn	Phe	Leu	Asp	Pro	Ser	Asp	Ala	Phe
	200					205					210				
Leu	Gly	Ala	Ser	Ser	Leu	Thr	Phe	Lys	Asn	Gly	Thr	Ala	Asn	Asp	Gly
215					220					225					230
Leu	Val	Gly	Thr	Cys	Ser	Ser	His	Leu	Gly	Met	Val	Ile	Arg	Asp	Asn
				235					240					245	
Tyr	Arg	Met	Asn	His	Leu	Asp	Glu	Val	Asn	Gln	Val	Phe	Gly	Leu	Thr
			250				255						260		
Ser	Leu	Phe	Glu	Thr	Ser	Pro	Val	Ser	Val	Tyr	Arg	Gln	His	Ala	Asn
		265					270					275			
Arg	Leu	Lys	Asn	Ala	Ser	Leu									
	280					285									

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1049 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION:85..1017

- (ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION:163..1017

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGATCCCCCG	GTTCTCCCGG	AAGGATTCGG	GCGATGGCTG	GCAGGACGCG	CCCCTCGGCC	60
CCATCAACCT	GAGATGAGAA	CAAC	ATG AAG AAG AAG	TCT CTG CTC	CCC CTC	111
			Met Lys Lys Lys	Ser Leu Leu	Pro Leu	
			-26 -25		-20	
GGC CTG GCC	ATC GGT CTC	GCC TCT CTC	GCT GCC AGC	CCT CTG ATC	CAG	159
Gly Leu Ala	Ile Gly Leu	Ala Ser Leu	Ala Ala Ser	Pro Leu Ile	Gln	
	-15		-10		-5	
GCC AGC ACC	TAC ACC CAG	ACC AAA TAC	CCC ATC GTG	CTG GCC CAC	GGC	207
Ala Ser Thr	Tyr Thr Gln	Thr Lys Tyr	Pro Ile Val	Leu Ala His	Gly	
	1	5	10		15	
ATG CTC GGC	TTC GAC AAC	ATC CTT GGG	GTC GAC TAC	TGG TTC GGC	ATT	255
Met Leu Gly	Phe Asp Asn	Ile Leu Gly	Val Asp Tyr	Trp Phe Gly	Ile	
	20		25		30	

CCC	AGC	GCC	TTG	CGC	CGT	GAC	GGT	GCC	CAG	GTC	TAC	GTC	ACC	GAA	GGC	303
Pro	Ser	Ala	Leu	Arg	Arg	Asp	Gly	Ala	Gln	Val	Tyr	Val	Thr	Glu	Gly	
			35					40					45			
AGC	CAG	TTG	GAC	ACC	TCG	GAA	GTC	CGC	GGC	GAG	CAG	TTG	CTG	CAA	CAG	351
Ser	Gln	Leu	Asp	Thr	Ser	Glu	Val	Arg	Gly	Glu	Gln	Leu	Gln	Gln		
		50					55					60				
GTG	GAG	GAA	ATC	GTC	GCC	CTC	AGC	GGC	CAG	CCC	AAG	GTC	AAC	CTG	ATC	399
Val	Glu	Glu	Ile	Val	Ala	Leu	Ser	Gly	Gln	Pro	Lys	Val	Asn	Leu	Ile	
	65					70				75						
GGC	CAC	AGC	CAC	GGC	GGG	CCG	ACC	ATC	CGC	TAC	GTC	GCC	GCC	GTA	CGT	447
Gly	His	Ser	His	Gly	Gly	Pro	Thr	Ile	Arg	Tyr	Val	Ala	Ala	Val	Arg	
	80				85					90					95	
CCC	GAC	CTG	ATC	GCT	TCC	GCC	ACC	AGC	GTC	GGC	GCC	CCG	CAC	AAG	GGT	495
Pro	Asp	Leu	Ile	Ala	Ser	Ala	Thr	Ser	Val	Gly	Ala	Pro	His	Lys	Gly	
				100					105					110		
TCG	GAC	ACC	GCC	GAC	TTC	CTG	CGC	CAG	ATC	CCA	CCG	GGT	TCG	GCC	GGC	543
Ser	Asp	Thr	Ala	Asp	Phe	Leu	Arg	Gln	Ile	Pro	Pro	Gly	Ser	Ala	Gly	
			115					120					125			
GAG	GCA	GTC	CTC	TCC	GGG	CTG	GTC	AAC	AGC	CTC	GGC	GCG	CTG	ATC	AGC	591
Glu	Ala	Val	Leu	Ser	Gly	Leu	Val	Asn	Ser	Leu	Gly	Ala	Leu	Ile	Ser	
		130					135					140				
TTC	CTT	TCC	AGC	GGC	GGC	ACC	GGT	ACG	CAG	AAT	TTA	CTG	GGC	TCG	CTG	639
Phe	Leu	Ser	Ser	Gly	Gly	Thr	Gly	Thr	Gln	Asn	Leu	Leu	Gly	Ser	Leu	
	145					150					155					
GAG	TCG	CTG	AAC	AGC	GAG	GGT	GCC	GCG	CGC	TTC	AAC	GCC	AAG	TAC	CCG	687
Glu	Ser	Leu	Asn	Ser	Glu	Gly	Ala	Ala	Arg	Phe	Asn	Ala	Lys	Tyr	Pro	
	160				165				170						175	
CAG	GGC	ATC	CCC	ACC	TCG	GCC	TGC	GGC	GAA	GGC	GCC	TAC	AAG	GTC	AAC	735
Gln	Gly	Ile	Pro	Thr	Ser	Ala	Cys	Gly	Glu	Gly	Ala	Tyr	Lys	Val	Asn	
				180					185					190		
GGC	GTG	AGC	TAT	TAC	TCC	TGG	AGC	GGT	TCC	TCG	CCG	CTG	ACC	AAC	TTC	783
Gly	Val	Ser	Tyr	Tyr	Ser	Trp	Ser	Gly	Ser	Ser	Pro	Leu	Thr	Asn	Phe	
			195					200					205			
CTC	GAT	CCG	AGC	GAC	GCC	TTC	CTC	GGC	GCC	TCG	TCG	CTG	ACC	TTC	AAG	831
Leu	Asp	Pro	Ser	Asp	Ala	Phe	Leu	Gly	Ala	Ser	Ser	Leu	Thr	Phe	Lys	
		210					215					220				
AAC	GGC	ACC	GCC	AAC	GAC	GGC	CTG	GTC	GGC	ACC	TGC	AGT	TCG	CAC	CTG	879
Asn	Gly	Thr	Ala	Asn	Asp	Gly	Leu	Val	Gly	Thr	Cys	Ser	Ser	His	Leu	
	225					230					235					
GGC	ATG	GTG	ATC	CGC	GAC	AAC	TAC	CGG	ATG	AAC	CAC	CTG	GAC	GAG	GTG	927
Gly	Met	Val	Ile	Arg	Asp	Asn	Tyr	Arg	Met	Asn	His	Leu	Asp	Glu	Val	
	240				245					250					255	
AAC	CAG	GTC	TTC	GGC	CTC	ACC	AGC	CTG	TTC	GAG	ACC	AGC	CCG	GTC	AGC	975
Asn	Gln	Val	Phe	Gly	Leu	Thr	Ser	Leu	Phe	Glu	Thr	Ser	Pro	Val	Ser	
				260					265					270		

GTC TAC CGC CAG CAC GCC AAC CGC CTG AAG AAC GCC AGC CTG 1017
Val Tyr Arg Gln His Ala Asn Arg Leu Lys Asn Ala Ser Leu
275 280 285

TAGGACCCCG GCCGGGGCCT CGGCCCGGGC CC 1049

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 311 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala
-26 -25 -20 -15
Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr
-10 -5 1 5
Lys Tyr Pro Ile Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile
10 15 20
Leu Gly Val Asp Tyr Trp Phe Gly Ile Pro Ser Ala Leu Arg Arg Asp
25 30 35
Gly Ala Gln Val Tyr Val Thr Glu Gly Ser Gln Leu Asp Thr Ser Glu
40 45 50
Val Arg Gly Glu Gln Leu Leu Gln Gln Val Glu Glu Ile Val Ala Leu
55 60 65 70
Ser Gly Gln Pro Lys Val Asn Leu Ile Gly His Ser His Gly Gly Pro
75 80 85
Thr Ile Arg Tyr Val Ala Ala Val Arg Pro Asp Leu Ile Ala Ser Ala
90 95 100
Thr Ser Val Gly Ala Pro His Lys Gly Ser Asp Thr Ala Asp Phe Leu
105 110 115
Arg Gln Ile Pro Pro Gly Ser Ala Gly Glu Ala Val Leu Ser Gly Leu
120 125 130
Val Asn Ser Leu Gly Ala Leu Ile Ser Phe Leu Ser Ser Gly Gly Thr
135 140 145 150
Gly Thr Gln Asn Leu Leu Gly Ser Leu Glu Ser Leu Asn Ser Glu Gly
155 160 165
Ala Ala Arg Phe Asn Ala Lys Tyr Pro Gln Gly Ile Pro Thr Ser Ala
170 175 180
Cys Gly Glu Gly Ala Tyr Lys Val Asn Gly Val Ser Tyr Tyr Ser Trp
185 190 195
Ser Gly Ser Ser Pro Leu Thr Asn Phe Leu Asp Pro Ser Asp Ala Phe
200 205 210

Leu Gly Ala Ser Ser Leu Thr Phe Lys Asn Gly Thr Ala Asn Asp Gly
 215 220 225 230
 Leu Val Gly Thr Cys Ser Ser His Leu Gly Met Val Ile Arg Asp Asn
 235 240 245
 Tyr Arg Met Asn His Leu Asp Glu Val Asn Gln Val Phe Gly Leu Thr
 250 255 260
 Ser Leu Phe Glu Thr Ser Pro Val Ser Val Tyr Arg Gln His Ala Asn
 265 270 275
 Arg Leu Lys Asn Ala Ser Leu
 280 285

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1050 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 85..1017

- (ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION: 163..1017

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GGATCCCCCG GTTCTCCCGG AAGGATTCGG GCGATGGCTG GCAGGACGCG CCCCTCGGCC 60
 CCATCAACCT GAGATGAGAA CAAC ATG AAG AAG AAG TCT CTG CTC CCC CTC 111
 Met Lys Lys Lys Ser Leu Leu Pro Leu
 -26 -25 -20
 GGC CTG GCC ATC GGT CTC GCC TCT CTC GCT GCC AGC CCT CTG ATC CAG 159
 Gly Leu Ala Ile Gly Leu Ala Ser Leu Ala Ala Ser Pro Leu Ile Gln
 -15 -10 -5
 GCC AGC ACC TAC ACC CAG ACC AAA TAC CCC ATC GTG CTG GCC CAC GGC 207
 Ala Ser Thr Tyr Thr Gln Thr Lys Tyr Pro Ile Val Leu Ala His Gly
 1 5 10 15
 ATG CTC GGC TTC GAC AAC ATC CTT GGG GTC GAC TAC TGG TTC GGC ATT 255
 Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp Phe Gly Ile
 20 25 30
 CCC AAC GCC TTG CGC CGT GAC GGT GCC CAG GTC TAC GTC ACC GAA GGC 303
 Pro Asn Ala Leu Arg Arg Asp Gly Ala Gln Val Tyr Val Thr Glu Gly
 35 40 45
 AGC CAG TTG GAC ACC TCG GAA GTC CGC GGC GAG CAG TTG CTG CAA CAG 351
 Ser Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gln Leu Leu Gln Gln
 50 55 60

GTG Val	GAG Glu	GAA Glu	ATC Ile	GTC Val	GCC Ala	CTC Leu	AGC Ser	GGC Gly	CAG Gln	CCC Pro	AAG Lys	GTC Val	AAC Asn	CTG Leu	ATC Ile	399
657075																
GGC Gly	CAC His	AGC Ser	CAC His	GGC Gly	GGG Gly	CCG Pro	ACC Thr	ATC Ile	CGC Arg	TAC Tyr	GTC Val	GCC Ala	GCC Ala	GTA Val	CGT Arg	447
80859095																
CCC Pro	GAC Asp	CTG Leu	ATC Ile	GCT Ala	TCC Ser	GCC Ala	ACC Thr	AGC Ser	GTC Val	GGC Gly	GCC Ala	CCG Pro	CAC His	AAG Lys	GGT Gly	495
100105																
TCG Ser	GAC Asp	ACC Thr	GCC Ala	GAC Asp	TTC Phe	CTG Leu	CGC Arg	CAG Gln	ATC Ile	CCA Pro	CCG Pro	GGT Gly	TCG Ser	GCC Ala	GGC Gly	543
115120125																
GAG Glu	GCA Ala	GTC Val	CTC Leu	TCC Ser	GGG Gly	CTG Leu	GTC Val	AAC Asn	AGC Ser	CTC Leu	GGC Gly	GCG Ala	CTG Leu	ATC Ile	AGC Ser	591
130135140																
TTC Phe	CTT Leu	TCC Ser	AGC Ser	GGC Gly	GGC Gly	ACC Thr	GGT Gly	ACG Thr	CAG Gln	AAT Asn	TTA Leu	CTG Leu	GGC Gly	TCG Ser	CTG Leu	639
145150155																
GAG Glu	TCG Ser	CTG Leu	AAC Asn	AGC Ser	GAG Glu	GGT Gly	GCC Ala	GCG Ala	CGC Arg	TTC Phe	AAC Asn	GCC Ala	AAG Lys	TAC Tyr	CCG Pro	687
160165170175																
CAG Gln	GGC Gly	ATC Ile	CCC Pro	ACC Thr	TCG Ser	GCC Ala	TGC Cys	GGC Gly	GAA Glu	GGC Gly	GCT Ala	TAC Tyr	AAG Lys	GTC Val	AAC Asn	735
180185190																
GGC Gly	GTG Val	AGC Ser	TAT Tyr	TAC Tyr	TCC Ser	TGG Trp	AGC Ser	GGT Gly	TCC Ser	TCG Ser	CCG Pro	CTG Leu	ACC Thr	AAC Asn	TTC Phe	783
195200205																
CTC Leu	GAT Asp	CCG Pro	AGC Ser	GAC Asp	GCC Ala	TTC Phe	CTC Leu	GGC Gly	GCC Ala	TCG Ser	TCG Ser	CTG Leu	ACC Thr	TTC Phe	AAG Lys	831
210215220																
AAC Asn	GGC Gly	ACC Thr	GCC Ala	AAC Asn	GAC Asp	GGC Gly	CTG Leu	GTC Val	GGC Gly	ACC Thr	TGC Cys	AGT Ser	TCG Ser	CAC His	CTG Leu	879
225230235																
GGC Gly	ATG Met	GTG Val	ATC Ile	CGC Arg	GAC Asp	AAC Asn	TAC Tyr	CGG Arg	ATG Met	AAC Asn	CAC His	CTG Leu	GAC Asp	GAG Glu	GTG Val	927
240245250255																
AAC Asn	CAG Gln	GTC Val	CTC Leu	GGC Gly	CTC Leu	ACC Thr	AGC Ser	CTG Leu	TTC Phe	GAG Glu	ACC Thr	AGC Ser	CCG Pro	GTC Val	AGC Ser	975
260265270275280285																
GTC Val	TAC Tyr	CGC Arg	CAG Gln	CAC His	GCC Ala	AAC Asn	CGC Arg	CTG Leu	AAG Lys	AAC Asn	GCC Ala	AGC Ser	CTG Leu			1017
275280285																
TAGGACCCCGGCCGGGGCCTCGGCCCGGGCCCGG																1050

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 311 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala
-26 -25 -20 -15
Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr
-10 -5 1 5
Lys Tyr Pro Ile Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile
10 15 20
Leu Gly Val Asp Tyr Trp Phe Gly Ile Pro Asn Ala Leu Arg Arg Asp
25 30 35
Gly Ala Gln Val Tyr Val Thr Glu Gly Ser Gln Leu Asp Thr Ser Glu
40 45 50
Val Arg Gly Glu Gln Leu Leu Gln Gln Val Glu Glu Ile Val Ala Leu
55 60 65 70
Ser Gly Gln Pro Lys Val Asn Leu Ile Gly His Ser His Gly Gly Pro
75 80 85
Thr Ile Arg Tyr Val Ala Ala Val Arg Pro Asp Leu Ile Ala Ser Ala
90 95 100
Thr Ser Val Gly Ala Pro His Lys Gly Ser Asp Thr Ala Asp Phe Leu
105 110 115
Arg Gln Ile Pro Pro Gly Ser Ala Gly Glu Ala Val Leu Ser Gly Leu
120 125 130
Val Asn Ser Leu Gly Ala Leu Ile Ser Phe Leu Ser Ser Gly Gly Thr
135 140 145 150
Gly Thr Gln Asn Leu Leu Gly Ser Leu Glu Ser Leu Asn Ser Glu Gly
155 160 165
Ala Ala Arg Phe Asn Ala Lys Tyr Pro Gln Gly Ile Pro Thr Ser Ala
170 175 180
Cys Gly Glu Gly Ala Tyr Lys Val Asn Gly Val Ser Tyr Tyr Ser Trp
185 190 195
Ser Gly Ser Ser Pro Leu Thr Asn Phe Leu Asp Pro Ser Asp Ala Phe
200 205 210
Leu Gly Ala Ser Ser Leu Thr Phe Lys Asn Gly Thr Ala Asn Asp Gly
215 220 225 230
Leu Val Gly Thr Cys Ser Ser His Leu Gly Met Val Ile Arg Asp Asn
235 240 245

Tyr Arg Met Asn His Leu Asp Glu Val Asn Gln Val Leu Gly Leu Thr
 250 255 260
 Ser Leu Phe Glu Thr Ser Pro Val Ser Val Tyr Arg Gln His Ala Asn
 265 270 275
 Arg Leu Lys Asn Ala Ser Leu
 280 285

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1049 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION:85..1017

- (ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION:163..1017

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GGATCCCCCG GTTCTCCCGG AAGGATTCGG GCGATGGCTG GCAGGACGCG CCCCTCGGCC 60
 CCATCAACCT GAGATGAGAA CAAC ATG AAG AAG AAG TCT CTG CTC CCC CTC 111
 Met Lys Lys Lys Ser Leu Leu Pro Leu
 -26 -25 -20
 GGC CTG GCC ATC GGT CTC GCC TCT CTC GCT GCC AGC CCT CTG ATC CAG 159
 Gly Leu Ala Ile Gly Leu Ala Ser Leu Ala Ala Ser Pro Leu Ile Gln
 -15 -10 -5
 GCC AGC ACC TAC ACC CAG ACC AAA TAC CCC ATC GTG CTG GCC CAC GGC 207
 Ala Ser Thr Tyr Thr Gln Thr Lys Tyr Pro Ile Val Leu Ala His Gly
 1 5 10 15
 ATG CTC GGC TTC GAC AAC ATC CTT GGG GTC GAC TAC TGG TTC GGC ATT 255
 Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp Phe Gly Ile
 20 25 30
 CCC AGC GCC TTG CGC CGT GAC GGT GCC CAG GTC TAC GTC ACC GAA GGC 303
 Pro Ser Ala Leu Arg Arg Asp Gly Ala Gln Val Tyr Val Thr Glu Gly
 35 40 45
 AGC CAG TTG GAC ACC TCG GAA GTC CGC GGC GAG CAG TTG CTG CAA CAG 351
 Ser Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gln Leu Leu Gln Gln
 50 55 60
 GTG GAG GAA ATC GTC GCC CTC AGC GGC CAG CCC AAG GTC AAC CTG ATC 399
 Val Glu Glu Ile Val Ala Leu Ser Gly Gln Pro Lys Val Asn Leu Ile
 65 70 75

GGC	CAC	AGC	CAC	GGC	GGG	CCG	ACC	ATC	CGC	TAC	GTC	GCC	GCC	GTA	CGT	447
Gly	His	Ser	His	Gly	Gly	Pro	Thr	Ile	Arg	Tyr	Val	Ala	Ala	Val	Arg	
80					85					90					95	
CCC	GAC	CTG	ATC	GCT	TCC	GCC	ACC	AGC	GTC	GGC	GCC	CCG	CAC	AGG	GGT	495
Pro	Asp	Leu	Ile	Ala	Ser	Ala	Thr	Ser	Val	Gly	Ala	Pro	His	Arg	Gly	
				100					105					110		
TCG	GAC	ACC	GCC	GAC	TTC	CTG	CGC	CAG	ATC	CCA	CCG	GGT	TCG	GCC	GGC	543
Ser	Asp	Thr	Ala	Asp	Phe	Leu	Arg	Gln	Ile	Pro	Pro	Gly	Ser	Ala	Gly	
			115					120					125			
GAG	GCA	GTC	CTC	TCC	GGG	CTG	GTC	AAC	AGC	CTC	GGC	GCG	CTG	ATC	AGC	591
Glu	Ala	Val	Leu	Ser	Gly	Leu	Val	Asn	Ser	Leu	Gly	Ala	Leu	Ile	Ser	
		130					135					140				
TTC	CTT	TCC	AGC	GGC	GGC	ACC	GGT	ACG	CAG	AAT	TTA	CTG	GGC	TCG	CTG	639
Phe	Leu	Ser	Ser	Gly	Gly	Thr	Gly	Thr	Gln	Asn	Leu	Leu	Gly	Ser	Leu	
	145					150					155					
GAG	TCG	CTG	AAC	AGT	GAG	GGT	GCC	GCG	CGC	TTC	AAC	GCC	AAG	TAC	CCG	687
Glu	Ser	Leu	Asn	Ser	Glu	Gly	Ala	Ala	Arg	Phe	Asn	Ala	Lys	Tyr	Pro	
160					165					170					175	
CAG	GGC	ATC	CCC	ACC	TCG	GCC	TGC	GGC	GAA	GGC	GCT	TAC	AAG	GTC	AAC	735
Gln	Gly	Ile	Pro	Thr	Ser	Ala	Cys	Gly	Glu	Gly	Ala	Tyr	Lys	Val	Asn	
				180				185						190		
GGC	GTG	AGC	TAT	TAC	TCC	TGG	AGC	GGT	TCC	TCG	CCG	CTG	ACC	AAC	TTC	783
Gly	Val	Ser	Tyr	Tyr	Ser	Trp	Ser	Gly	Ser	Ser	Pro	Leu	Thr	Asn	Phe	
			195					200					205			
CTC	GAT	CCG	AGC	GAC	GCC	TTC	CTC	GGC	GCC	TCG	TCG	CTG	ACC	TTC	AAG	831
Leu	Asp	Pro	Ser	Asp	Ala	Phe	Leu	Gly	Ala	Ser	Ser	Leu	Thr	Phe	Lys	
		210					215					220				
AAC	GGC	ACC	GCC	AAC	GAC	GGC	CTG	GTC	GGC	ACC	TGC	AGT	TCG	CAC	CTG	879
Asn	Gly	Thr	Ala	Asn	Asp	Gly	Leu	Val	Gly	Thr	Cys	Ser	Ser	His	Leu	
	225					230					235					
GGC	ATG	GTG	ATC	CGC	GAC	AAC	TAC	CGG	ATG	AAC	CAC	CTG	GAC	GAG	GTG	927
Gly	Met	Val	Ile	Arg	Asp	Asn	Tyr	Arg	Met	Asn	His	Leu	Asp	Glu	Val	
240					245					250				255		
AAC	CAG	GTC	CTC	GGC	CTC	ACC	AGC	CTG	TTC	GAG	ACC	AGC	CCG	GTC	AGC	975
Asn	Gln	Val	Leu	Gly	Leu	Thr	Ser	Leu	Phe	Glu	Thr	Ser	Pro	Val	Ser	
				260				265						270		
GTC	TAC	CGC	CAG	CAC	GCC	AAC	CGC	CTG	AAG	AAC	GCC	AGC	CTG			1017
Val	Tyr	Arg	Gln	His	Ala	Asn	Arg	Leu	Lys	Asn	Ala	Ser	Leu			
			275					280					285			
TAGGACCCCCG GCCGGGGCCT CGGCCCGGGC CC																1049

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 311 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

```

Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala
-26 -25                -20                -15

Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr
-10                -5                1                5

Lys Tyr Pro Ile Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile
                10                15                20

Leu Gly Val Asp Tyr Trp Phe Gly Ile Pro Ser Ala Leu Arg Arg Asp
                25                30                35

Gly Ala Gln Val Tyr Val Thr Glu Gly Ser Gln Leu Asp Thr Ser Glu
                40                45                50

Val Arg Gly Glu Gln Leu Leu Gln Gln Val Glu Glu Ile Val Ala Leu
                55                60                65                70

Ser Gly Gln Pro Lys Val Asn Leu Ile Gly His Ser His Gly Gly Pro
                75                80                85

Thr Ile Arg Tyr Val Ala Ala Val Arg Pro Asp Leu Ile Ala Ser Ala
                90                95                100

Thr Ser Val Gly Ala Pro His Arg Gly Ser Asp Thr Ala Asp Phe Leu
                105                110                115

Arg Gln Ile Pro Pro Gly Ser Ala Gly Glu Ala Val Leu Ser Gly Leu
                120                125                130

Val Asn Ser Leu Gly Ala Leu Ile Ser Phe Leu Ser Ser Gly Gly Thr
                135                140                145                150

Gly Thr Gln Asn Leu Leu Gly Ser Leu Glu Ser Leu Asn Ser Glu Gly
                155                160                165

Ala Ala Arg Phe Asn Ala Lys Tyr Pro Gln Gly Ile Pro Thr Ser Ala
                170                175                180

Cys Gly Glu Gly Ala Tyr Lys Val Asn Gly Val Ser Tyr Tyr Ser Trp
                185                190                195

Ser Gly Ser Ser Pro Leu Thr Asn Phe Leu Asp Pro Ser Asp Ala Phe
                200                205                210

Leu Gly Ala Ser Ser Leu Thr Phe Lys Asn Gly Thr Ala Asn Asp Gly
                215                220                225                230

Leu Val Gly Thr Cys Ser Ser His Leu Gly Met Val Ile Arg Asp Asn
                235                240                245

Tyr Arg Met Asn His Leu Asp Glu Val Asn Gln Val Leu Gly Leu Thr
                250                255                260

Ser Leu Phe Glu Thr Ser Pro Val Ser Val Tyr Arg Gln His Ala Asn
                265                270                275

Arg Leu Lys Asn Ala Ser Leu
                280                285

```

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1049 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION:85..1017

- (ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION:163..1017

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

```

GGATCCCCCG GTTCTCCCGG AAGGATTCGG GCGATGGCTG GCAGGACGCG CCCCTCGGCC      60
CCATCAACCT GAGATGAGAA CAAC ATG AAG AAG AAG TCT CTG CTC CCC CTC      111
               Met Lys Lys Lys Ser Leu Leu Pro Leu
               -26 -25                               -20

GGC CTG GCC ATC GGT CTC GCC TCT CTC GCT GCC AGC CCT CTG ATC CAG      159
Gly Leu Ala Ile Gly Leu Ala Ser Leu Ala Ala Ser Pro Leu Ile Gln
               -15                               -10                               -5

GCC AGC ACC TAC ACC CAG ACC AAA TAC CCC ATC GTG CTG GCC CAC GGC      207
Ala Ser Thr Tyr Thr Gln Thr Lys Tyr Pro Ile Val Leu Ala His Gly
               1                               5                               10                               15

ATG CTC GGC TTC GAC AAC ATC CTT GGG GTC GAC TAC TGG TTC GGC ATT      255
Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp Phe Gly Ile
               20                               25                               30

CCC AGC GCC TTG CGC CGT GAC GGT GCC CAG GTC TAC GTC ACC GAA GGC      303
Pro Ser Ala Leu Arg Arg Asp Gly Ala Gln Val Tyr Val Thr Glu Gly
               35                               40                               45

AGC CAG TTG GAC ACC TCG GAA GTC CGC GGC GAG CAG TTG CTG CAA CAG      351
Ser Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gln Leu Leu Gln Gln
               50                               55                               60

GTG GAG GAA ATC GTC GCC CTC AGC GGC CAG CCC AAG GTC AAC CTG ATC      399
Val Glu Glu Ile Val Ala Leu Ser Gly Gln Pro Lys Val Asn Leu Ile
               65                               70                               75

GGC CAC AGC CAC GGC GGG CCG ACC ATC CGC TAC GTC GCC GCC GTA CGT      447
Gly His Ser His Gly Gly Pro Thr Ile Arg Tyr Val Ala Ala Val Arg
               80                               85                               90                               95

CCC GAC CTG ATC GCT TCC GCC ACC AGC GTC GGC GCC CCG CAC AAG GGT      495
Pro Asp Leu Ile Ala Ser Ala Thr Ser Val Gly Ala Pro His Lys Gly
               100                               105                               110

```

TCG GAC ACC GCC GAC TTC CTG CGC CAG ATC CCA CCG GGT TCG GCC GGC	543
Ser Asp Thr Ala Asp Phe Leu Arg Gln Ile Pro Pro Gly Ser Ala Gly	
115 120 125	
GAG GCA GTC CTC TCC GGG CTG GTC AAC AGC CTC GGC GCG CTG ATC AGC	591
Glu Ala Val Leu Ser Gly Leu Val Asn Ser Leu Gly Ala Leu Ile Ser	
130 135 140	
TTC CTT TCC AGC GGC GGC ATC GGT ACG CAG AAT TTT CTG GGC TCG CTG	639
Phe Leu Ser Ser Gly Gly Ile Gly Thr Gln Asn Phe Leu Gly Ser Leu	
145 150 155	
GAG TCG CTG AAC AGC GAG GGT GCC GCG CGC TTC AAC GCC AAG TAC CCG	687
Glu Ser Leu Asn Ser Glu Gly Ala Ala Arg Phe Asn Ala Lys Tyr Pro	
160 165 170 175	
CAG GGC ATC CCC ACC TCG GCC TGC GGC GAA GGC GCC TAC AAG GTC AAC	735
Gln Gly Ile Pro Thr Ser Ala Cys Gly Glu Gly Ala Tyr Lys Val Asn	
180 185 190	
GGC GTG AGC TAT TAC TCC TGG AGC GGT TCC TCG CCG CTG ACC AAC TTC	783
Gly Val Ser Tyr Tyr Ser Trp Ser Gly Ser Ser Pro Leu Thr Asn Phe	
195 200 205	
CTC GAT CCG AGC GAC GCC TTC CTC GGC GCC TCG TCG CTG ACC TTC AAG	831
Leu Asp Pro Ser Asp Ala Phe Leu Gly Ala Ser Ser Leu Thr Phe Lys	
210 215 220	
AAC GGC ACC GCC AAC GAC GGC CTG GTC GGC ACC TGC AGT TCG CAC CTG	879
Asn Gly Thr Ala Asn Asp Gly Leu Val Gly Thr Cys Ser Ser His Leu	
225 230 235	
GGC ATG GTG ATC CGC GAC AAC TAC CGG ATG AAC CAC CTG GAC GAG GTG	927
Gly Met Val Ile Arg Asp Asn Tyr Arg Met Asn His Leu Asp Glu Val	
240 245 250 255	
AAC CAG GTC TTC GGC CTC ACC AGC CTG TTC GAG ACC AGC CCG GTC AGC	975
Asn Gln Val Phe Gly Leu Thr Ser Leu Phe Glu Thr Ser Pro Val Ser	
260 265 270	
GTC TAC CGC CAG CAC GCC AAC CGC CTG AAG AAC GCC AGC CTG	1017
Val Tyr Arg Gln His Ala Asn Arg Leu Lys Asn Ala Ser Leu	
275 280 285	
TAGGACCCCG GCCGGGGCCT CGGCCCGGGC CC	1049

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 311 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala	
-26 -25 -20 -15	
Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr	
-10 -5 1 5	

Lys	Tyr	Pro	Ile 10	Val	Leu	Ala	His	Gly 15	Met	Leu	Gly	Phe	Asp 20	Asn	Ile
Leu	Gly	Val 25	Asp	Tyr	Trp	Phe	Gly 30	Ile	Pro	Ser	Ala	Leu 35	Arg	Arg	Asp
Gly	Ala 40	Gln	Val	Tyr	Val	Thr 45	Glu	Gly	Ser	Gln	Leu 50	Asp	Thr	Ser	Glu
Val 55	Arg	Gly	Glu	Gln	Leu 60	Leu	Gln	Gln	Val	Glu 65	Glu	Ile	Val	Ala	Leu 70
Ser	Gly	Gln	Pro	Lys 75	Val	Asn	Leu	Ile	Gly 80	His	Ser	His	Gly	Gly 85	Pro
Thr	Ile	Arg	Tyr 90	Val	Ala	Ala	Val	Arg 95	Pro	Asp	Leu	Ile	Ala 100	Ser	Ala
Thr	Ser	Val 105	Gly	Ala	Pro	His	Lys 110	Gly	Ser	Asp	Thr	Ala 115	Asp	Phe	Leu
Arg	Gln 120	Ile	Pro	Pro	Gly	Ser 125	Ala	Gly	Glu	Ala	Val 130	Leu	Ser	Gly	Leu
Val 135	Asn	Ser	Leu	Gly	Ala 140	Leu	Ile	Ser	Phe	Leu 145	Ser	Ser	Gly	Gly	Ile 150
Gly	Thr	Gln	Asn	Phe 155	Leu	Gly	Ser	Leu	Glu 160	Ser	Leu	Asn	Ser	Glu 165	Gly
Ala	Ala	Arg	Phe 170	Asn	Ala	Lys	Tyr	Pro 175	Gln	Gly	Ile	Pro	Thr 180	Ser	Ala
Cys	Gly	Glu 185	Gly	Ala	Tyr	Lys	Val 190	Asn	Gly	Val	Ser	Tyr 195	Tyr	Ser	Trp
Ser	Gly 200	Ser	Ser	Pro	Leu	Thr 205	Asn	Phe	Leu	Asp	Pro 210	Ser	Asp	Ala	Phe
Leu 215	Gly	Ala	Ser	Ser	Leu 220	Thr	Phe	Lys	Asn	Gly 225	Thr	Ala	Asn	Asp	Gly 230
Leu	Val	Gly	Thr	Cys 235	Ser	Ser	His	Leu	Gly 240	Met	Val	Ile	Arg	Asp 245	Asn
Tyr	Arg	Met	Asn 250	His	Leu	Asp	Glu	Val 255	Asn	Gln	Val	Phe	Gly 260	Leu	Thr
Ser	Leu	Phe 265	Glu	Thr	Ser	Pro	Val 270	Ser	Val	Tyr	Arg	Gln 275	His	Ala	Asn
Arg	Leu 280	Lys	Asn	Ala	Ser	Leu 285									

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEOUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

GCGCAATTAA CCCTCACTAA AGGGAACAAA

30

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

GGTACGCAGA ATNNNCTGGG CTCGC

25

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GCGTAATACG ACTCACTATA GGGCGAA

27